AGU ADVANCING EARTH AND SPACE SCIENCES

Geophysical Research Letters

RESEARCH LETTER
10.1029/2024GL114298

Toward Artificial General Intelligence in Hydrogeological Modeling With an Integrated Latent Diffusion Framework

Chuanjun Zhan[1,2], Zhenxue Dai[1,3], Jiu Jimmy Jiao[2], Mohamad Reza Soltanian[4], Huichao Yin[5], and Kenneth C. Carroll[5]

[1]School of Environmental and Municipal Engineering, Qingdao University of Technology, Qingdao, China, [2]Department of Earth Sciences, The University of Hong Kong, Hong Kong, China, [3]College of Construction Engineering, Jilin University, Changchun, China, [4]Departments of Geosciences and Environmental Engineering, University of Cincinnati, Cincinnati, OH, USA, [5]Department of Plant & Environmental Science, New Mexico State University, Las Cruces, NM, USA

Key Points:

- Latent Diffusion Framework enables hydrogeological modeling across multiple tasks
- The integrated framework achieves diverse hydrogeological objectives without task-specific designs
- The multi-task capabilities highlight the potential for Artificial General Intelligence in hydrogeological modeling

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Z. Dai and J. J. Jiao,
dzx@jlu.edu.cn;
jjiao@hku.hk

Citation:

Zhan, C., Dai, Z., Jiao, J. J., Soltanian, M. R., Yin, H., & Carroll, K. C. (2025). Toward artificial general intelligence in hydrogeological modeling with an integrated latent diffusion framework. *Geophysical Research Letters*, *52*, e2024GL114298. https://doi.org/10.1029/2024GL114298

Received 17 DEC 2024
Accepted 18 JAN 2025

Author Contributions:

Conceptualization: Chuanjun Zhan, Zhenxue Dai, Mohamad Reza Soltanian, Kenneth C. Carroll
Funding acquisition: Zhenxue Dai, Jiu Jimmy Jiao
Methodology: Chuanjun Zhan
Project administration: Zhenxue Dai, Jiu Jimmy Jiao
Supervision: Zhenxue Dai, Jiu Jimmy Jiao
Validation: Chuanjun Zhan
Visualization: Chuanjun Zhan
Writing – original draft: Chuanjun Zhan
Writing – review & editing: Chuanjun Zhan, Zhenxue Dai, Jiu Jimmy Jiao, Mohamad Reza Soltanian, Huichao Yin, Kenneth C. Carroll

Abstract Deep learning models have been extensively applied to various aspects of hydrogeological modeling. However, traditional approaches often rely on separate task-specific models, resulting in time-consuming selection and tuning processes. This study develops an integrated Latent Diffusion Model (LDM) framework to address four key hydrogeological modeling tasks: aquifer heterogeneity structure generation, surrogate modeling for flow and transport, and direct inversion of aquifer heterogeneity structure. Using a consistent architecture and hyperparameters, the LDM demonstrates robust multi-task processing capabilities, accurately capturing aquifer heterogeneity, enabling rapid predictions of hydraulic head and solute transport, and efficiently performing direct inversion without iterative simulations. By integrating multiple tasks within a single framework, LDM eliminates the need for task-specific models or extensive parameter optimization, offering an efficient and adaptive general solution for deep learning-based hydrogeological modeling. Its generalization across diverse objectives underscores its potential as a cornerstone for advancing Artificial General Intelligence in hydrogeological modeling.

Plain Language Summary Hydrogeological modeling is critical for water resource management and environmental protection, but their complex and variable nature makes accurate modeling challenging. Traditional approaches rely on separate models for tasks such as simulating water flow, predicting solute transport, or identifying subsurface structures, and each task requires significant time and effort for setup and adjustment. This study introduces an Latent Diffusion Model (LDM) framework that efficiently performs multiple hydrogeological modeling tasks. The LDM generates detailed representations of subsurface heterogeneity, efficiently predicts flow and mass transport, and directly identifies underground structures from observed data without iterative simulations. By integrating these capabilities into a single framework, the LDM streamlines workflows, reduces computational demands, and improves adaptability across diverse scenarios. This innovative approach not only enhances modeling efficiency but also lays the foundation for intelligent systems to address complex environmental challenges.

1. Introduction

Artificial General Intelligence (AGI) refers to a system that can understand, learn, and apply knowledge across a wide range of tasks, exhibiting cognitive abilities similar to those of humans. While significant progress has been made in narrow AI, true AGI remains an elusive goal. Key characteristics include the ability for autonomous reasoning, transfer learning, and domain-independent problem-solving (Fei et al., 2022; Kuusi & Heinonen, 2022). Among various capabilities required for AGI development, multi-task processing represents an essential component, particularly in domains like hydrogeological modeling where integrated approaches are needed to address complex, interconnected problems. The development of AGI in hydrogeological modeling could revolutionize how we approach subsurface characterization, enabling systems that can seamlessly transition between different modeling tasks, adapt to new scenarios without extensive retraining, and integrate diverse data sources in ways that mimic human expertise. Such capabilities would be particularly valuable for addressing challenges like real-time aquifer management, complex contamination scenarios, and climate change impacts on groundwater systems. This transformative vision holds significant potential for numerous fields, including environmental and earth sciences, where the complexity and variability of systems (e.g., groundwater flow and subsurface contaminant migration) necessitate more integrated and flexible modeling approaches.

Accurate modeling of groundwater flow, solute transport, and aquifer heterogeneity is essential for effective water resource management and environmental protection (Harp et al., 2008; Kitanidis, 2015; Rajaram & Gelhar, 1995; Rizzo & de Barros, 2019). However, the inherent complexity of subsurface environments, characterized by multiscale heterogeneity and limited observational data, presents significant challenges for traditional modeling approaches (Carrera et al., 2005; Jankovic et al., 2017; Scheibe et al., 2015; Song et al., 2019; Zhu & Yeh, 2005). Existing methods often rely on a series of task-specific models, each optimized for a single objective, such as predicting contaminant transport or describing aquifer structures. While these models have demonstrated utility, they are time-consuming, computationally expensive, and require extensive calibration and optimization for each task (Sbai, 2020; Zhan et al., 2023). Moreover, the need to integrate diverse data sources and types, such as borehole measurements, hydraulic head distributions, and solute concentrations, further complicates traditional modeling workflows.

Recent advances in deep learning (DL) have shown promise in addressing some of these challenges by offering more flexible, data-driven approaches to subsurface modeling (Ershadnia et al., 2024; Kang et al., 2021; Mo et al., 2020; J. Zhang et al., 2024). However, achieving different objectives often require distinct model architectures, resulting in a proliferation of task-specific designs. Previous studies frequently employed separate DL models tailored to individual tasks, such as flow prediction or aquifer structure characterization (Cui et al., 2024; Moeini et al., 2024; Zhan, Dai, Soltanian, & Zhang, 2022), without a unified approach. The selection and tuning of models is a time-intensive process, which may offset the efficiency gains promised by DL methods.

From a lifecycle perspective, the reliance on multiple specialized models may limit the advantages of DL approaches over traditional computational methods, thereby constraining their broader adoption in hydrogeology to overcome these limitations, the multi-task processing capabilities of AGI present a compelling alternative. AGI enables a single model to handle diverse objectives without requiring task-specific architectures or extensive parameter adjustments (Feng et al., 2024). This integrated approach enhances the overall efficiency of subsurface modeling and represents a significant step toward realizing the vision of AGI in hydrogeological applications.

Latent Diffusion Model (LDM), as a next-generation generative artificial intelligence model, offer a promising solution to this challenge. LDM have been widely applied in various image generation tasks (Rombach et al., 2022). Popular models such as Stable Diffusion and DALL·E 3, which power advanced tools such as ChatGPT, are either based on or incorporate LDM (Bengesi et al., 2024). These models excel at a variety of image processing tasks, including generating images conditioned on text or other inputs, modifying specific regions, filling missing areas, and repairing damaged or aged images (Yang et al., 2023). Compared to earlier generative models, such as Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs), LDM offer more stable training, enhanced control over conditional generation, and the versatility to handle multiple tasks (Lee et al., 2024). Moreover, unlike other diffusion models such as Denoising Diffusion Probabilistic Models (DDPMs), LDM operate in latent space, significantly reducing training time and computational resource requirements (Blattmann et al., 2023).

Inspired by the success of LDM in image generation, their robust training processes, strong conditional control, and multi-task learning capabilities make them highly suitable for addressing various challenges in hydrogeological modeling (Di Federico & Durlofsky, 2024; X. Zhang, Jiang, Wei, et al., 2024). This study investigates the potential of LDM in advancing multi-task processing capabilities for hydrogeological modeling applications. By leveraging the multi-task capabilities of LDM, we aim to address four critical tasks in hydrogeological modeling: generating single-scale aquifer heterogeneity, generating multiscale heterogeneity, constructing surrogate models for groundwater flow and solute transport, and directly identifying aquifer heterogeneity structure from observational data. These tasks are performed within an integrated LDM framework, leading to the development of a versatile hydrogeological modeling framework that facilitates multiscale aquifer characterization and eliminates the need for multiple specialized models. This single-model approach provides a powerful alternative to traditional hydrogeological modeling, enhancing computational efficiency and adaptability across diverse scenarios. Furthermore, this contributes to exploring the multi-task processing capabilities required for future AGI models in hydrogeology.

2. Problem Statement

This section outlines four key tasks designed to address challenges in hydrogeological modeling using a LDM framework:

2.1. Single-Scale Aquifer Heterogeneity Structure Generation

This task evaluates the ability of LDM to represent single-scale aquifer heterogeneity using low-dimensional latent random variables. Acting as a parameterization model, it facilitates uncertainty analysis of heterogeneous aquifer structures (Laloy et al., 2018). For this, 10,000 two-dimensional aquifer facies structures were generated as training samples using indicator kriging simulation based on transition probabilities (details in Supporting Information S1, Text S1 in Supporting Information S1). The data set was built using data from five hypothetical boreholes (Figure S2 in Supporting Information S1). Each facies structure measures 80 m × 40 m and is discretized into 3,200 uniform grids, comprising two facies: high-permeability and low-permeability zones (see Figure 2). Such type of heterogeneity, that is, binary, has been shown to control flow and transport process in the subsurface environment (Dai et al., 2005; Gershenzon et al., 2015).

2.2. Multiscale Aquifer Heterogeneity Structure Generation

Building upon Task 1, this task investigates the performance of LDM in generating multiscale aquifer heterogeneity by incorporating permeability heterogeneity within facies. This ensures that the permeability distribution is constrained by the large-scale facies structure, with permeability values within each facies adhering to specified geological statistical distributions. The training data set comprises the 10,000 large-scale facies structures from Task 1, combined with 10,000 small-scale permeability fields generated using the Karhunen–Loève expansion (KLE) (Yang et al., 2004). Two different Gaussian distributions were considered for the permeability fields within the two lithologies. Specifically, the permeability fields corresponding to the high- and low-permeability facies have mean log-permeability values of 4 and 0, respectively, both facies share the same variance of the log-transformed permeability, given by $\sigma^2_{\mathrm{Log}\ (K)} = 0.5$,the correlation lengths differ between facies, with the high-permeability facies having a correlation length of $\lambda_H = 15$ and the low-permeability facies having $\lambda_L = 10$. Details on determining multiscale heterogeneity priors are provided in Text S5 in Supporting Information S1. While we demonstrated the framework using binary facies, it can be extended to multiple facies by simply training with multi-facies samples generated from GEOST and KLE, without modifying the model architecture or training process.

2.3. Groundwater Flow and Solute Transport Surrogate Modeling

This task explores the use of LDM to construct surrogate models for groundwater flow and solute transport in heterogeneous aquifers. These surrogate models rapidly map aquifer facies structures to observed data, such as hydraulic heads and solute concentrations, accelerating iterative processes in data assimilation. Using the 10,000 facies structures from Task 1, tracer tests were simulated over two months using TOUGHREACT. The model domain featured constant hydraulic head boundaries on the left and right, creating a hydraulic gradient of 5×10^{-3}, while other boundaries were impermeable. A tracer was injected at a rate of 0.04 kg/s across the full thickness near the left boundary (See Figure S2 in Supporting Information S1). Hydraulic head and solute concentration fields were recorded at five time points during injection. This generated a training data set consisting of facies structures paired with their corresponding hydraulic head and solute concentration fields.

2.4. Direct Inversion of Aquifer Heterogeneity Structure

This task assesses the ability of LDM to directly invert aquifer heterogeneity from observational data, such as hydraulic head and solute concentration fields, without iterative simulation or traditional data assimilation methods. The same training data from Task 3 were used, but the model was applied in reverse, predicting large-scale facies structures from observational data. Details on this process are provided in Section 3.

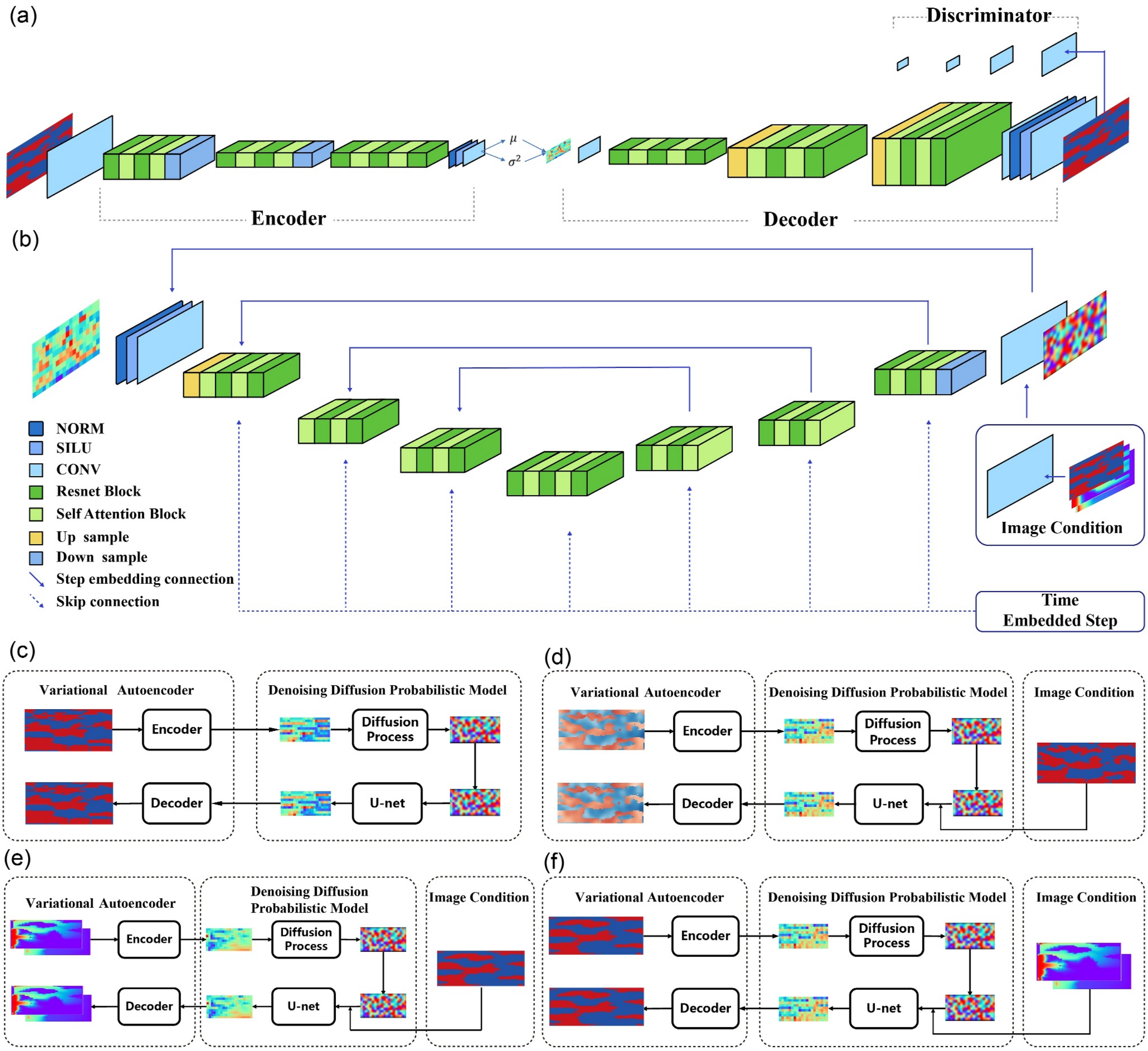

Figure 1. The Latent Diffusion Model framework and task implementation: (a) Variational Autoencoder, (b) U-Net, and schematic diagrams for (c) Single-scale heterogeneity generation, (d) Multiscale heterogeneity generation, (e) Surrogate modeling, and (f) Direct inversion.

3. Method

This section describes the LDM framework implemented to address four key tasks in aquifer heterogeneity characterization. The methodology includes an overview of the LDM architecture, task-specific adaptations, and implementation steps.

3.1. Overview of Latent Diffusion Models Framework

The LDM framework comprises three main components: an encoder, a denoising diffusion probabilistic model (DDPM), and a decoder (See Figure 1a). The encoder compresses high-dimensional data into a compact latent space, while the decoder maps the refined latent variables back to the high-dimensional data space. DDPM, the core component, includes forward diffusion and reverse denoising phases. The forward phase progressively adds Gaussian noise to latent variables using a linear noise scheduler (Figure S3 in Supporting Information S1), while the reverse phase employs a U-Net neural network to remove noise and reconstruct the data. A conditioning mechanism enables sample generation based on external conditions during denoising. The framework utilizes residual network modules and attention mechanisms in its encoder, decoder, and U-Net components. To explore general performance across different tasks, the structure remains consistent, maintaining the same number of residual and attention modules, convolutional channel dimensions, noise levels, timesteps, and iterations.

Additionally, LDM's conditional control mechanism enables sample generation based on external conditions (e.g., images or other features) during denoising. In this study, images representing facies, hydraulic head, or concentration distributions serve as conditioning information. The conditioning process involves resizing the control image to match the latent variable's dimensions and directly stacking them with the latent variables at each denoising step during each denoising step (see Figure 1b), the architectural details and process description of the LDM are given in Supporting Information S1: Texts S3 and S4 in Supporting Information S1. By concatenating the conditioned image with the latent variable at each step, the model generates samples that meet the specified conditions. The final conditioned latent variable is then fed into the pre-trained decoder to predict facies structures or distributions of hydraulic head and concentration under specified conditions.

In this study, residual network modules and attention mechanisms are utilized to construct the encoder, decoder, and the U-Net within the DDPM of the LDM model, as shown in Figures 1a and 1b. To explore the general performance of the LDM model across different tasks, the structure of the encoder, decoder, and DDPM remains consistent across tasks. This includes maintaining the same number of residual and attention modules, convolutional channel dimensions, initial and final noise levels, total timesteps, and number of iterations. Detailed descriptions of the specific module structures and network configurations are provided in the Supporting Information S1.

3.2. Implementation for Specific Tasks

The LDM framework is capable of performing different tasks by utilizing the training data sets described in Section 2. By training the model on these data sets, various functions can be achieved. The specific implementation for each task is as follows:

3.2.1. Single-Scale Aquifer Heterogeneity Structure Generation

As shown in Figure 1c, the training process begins by encoding and decoding 10,000 aquifer facies-structure samples. After training, the encoder converts these 10,000 facies-structure samples into 10,000 corresponding latent variable samples. In this study, the size of the latent variables for all tasks is set to be one-fourth of the original data. The DDPM model is then trained using these latent variables, where Gaussian noise is iteratively added in the forward diffusion process, followed by the reverse diffusion process that gradually removes the noise. The U-Net network is used in the reverse process to minimize the mean squared error (MSE) between predicted and actual latent variables.

Importantly, no conditional constraints are applied when generating single-scale aquifer heterogeneity structure. After the model is trained, a random Gaussian noise vector is generated and passed through the U-Net model, which applies the reverse diffusion process to generate a new latent variable. This latent variable is then decoded to produce the corresponding aquifer facies structure. Similarly, the same approach can be used to generate permeability fields, with the only difference being that the training data for the encoder and decoder would be replaced with permeability fields instead of facies structures.

3.2.2. Multiscale Aquifer Heterogeneity Structure Generation

As shown in Figure 1d, the training process begins by using 10,000 permeability fields as the training data to train the encoder and decoder. After training, the encoder converts the permeability fields into corresponding latent variables. During the DDPM training phase, the facies-structure samples corresponding to the permeability fields are used as conditional constraints. These are integrated into the reverse diffusion process through a convolutional layer, which controls the formation of the final latent random variable. Once the DDPM is trained, a random Gaussian noise vector is generated, and the corresponding facies structure is given as a condition. The latent variable is generated and then passed through the decoder to generate the permeability field, constrained by the facies distribution in space.

3.2.3. Groundwater Flow and Solute Transport Surrogate Modeling

As shown in Figure 1e, during surrogate model construction, the encoder and decoder are trained using hydraulic head and concentration distribution fields as training data. The model learns to convert these fields into latent variables. The DDPM reverse diffusion process is guided by the aquifer facies structure, generating the corresponding latent variables, which are then decoded into the corresponding hydraulic head and concentration distribution fields. This enables a rapid mapping from aquifer facies structures to hydraulic head and concentration fields.

3.2.4. Direct Inversion of Aquifer Heterogeneity Structure

As shown in Figure 1f, the direct inversion task is the reverse of the surrogate model mapping. In this task, hydraulic head and concentration fields are used as inputs to predict the aquifer facies structure. The training process involves learning and training facies structure samples using a VAE, encoding them into low-dimensional latent variables. During the DDPM training phase, hydraulic head and concentration fields serve as constraints, guiding the generation of latent variables. After training, for a given hydraulic head and concentration field, a corresponding facies structure is generated by applying the reverse diffusion process to a random Gaussian noise input and passing it through the decoder.

4. Results and Discussion

This section presents the results of applying the LDM framework to the four tasks outlined in Section 2. All tasks utilized the same VAE and DDPM architecture without any modifications to the network layers, channels, loss functions, learning rates, or batch sizes. This intentional design choice aimed to investigate the generalization capability of the LDM framework and evaluate its potential as a unified tool for multi-task hydrogeological modeling, contributing to the advancement of AGI in hydrogeology.

4.1. Single-Scale Aquifer Heterogeneity Generation

The LDM framework successfully generated single-scale aquifer facies structures with statistical and spatial properties closely matching those of the training samples. Figure 2 provides a detailed comparison between the generated and training structures. The reconstructed facies probability distribution (Figure 2d) aligns almost perfectly with the training data (Figure 2a), indicating that the model effectively captures the spatial distribution of facies in aquifer structures. Additionally, the randomly generated facies structures (Figures 2e and 2f) exhibit spatial connectivity and distribution patterns consistent with the training samples (Figures 2b and 2c).

From a geological statistical perspective, the LDM framework preserved key volumetric and geometric features of the training data. For example, as shown in Figures 2g–2i, the generated structures demonstrate a similar volume proportion of low-permeability facies, accurately reflecting the overall facies distribution. The horizontal and vertical lengths of low-permeability facies closely match those of the training structures, highlighting the model's ability to capture directional continuity and anisotropy in aquifer systems.

This task emphasizes the LDM's ability to generate realistic single-scale aquifer heterogeneity without requiring explicit conditional constraints. Although this study focused on binary facies structures, the framework can easily be adapted to generate other single-scale heterogeneity representations, such as permeability fields, by replacing the VAE training samples. Compared to traditional parameterization models, such as GANs or VAEs, the LDM offers improved training stability and generation performance (Lee et al., 2024; X. Zhang, Jiang, Zheng, et al., 2024), making it a robust tool for stochastic generation and uncertainty analysis.

4.2. Multiscale Aquifer Heterogeneity Generation

Building on single-scale generation, the LDM framework demonstrated strong performance in generating multiscale aquifer heterogeneity by incorporating permeability fields within facies distributions. Figure 2j compares the permeability distributions of the training samples with those generated by the model, while Figures 2k and 2l display examples of generated permeability fields conditioned on two randomly selected facies structures. These results indicate that the permeability fields generated by the LDM adhere to the geological constraints defined by the facies, with permeability values conditioned on the spatial extent and type of facies. Each facies type exhibits distinct permeability distributions, and the generated permeability fields have a bimodal permeability coefficient

Geophysical Research Letters

10.1029/2024GL114298

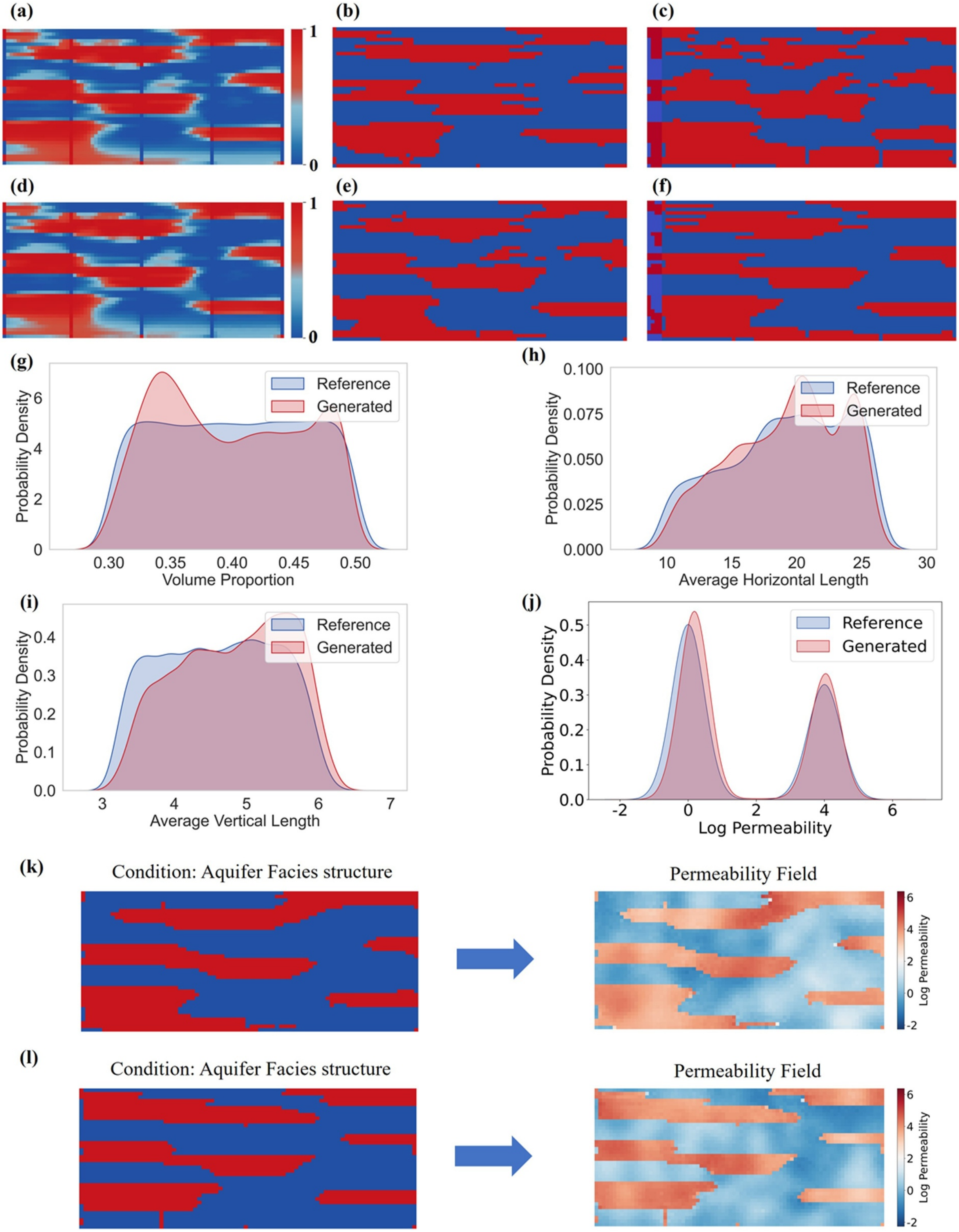

Figure 2. Comparison of aquifer facies and permeability fields generated by Latent Diffusion Model (LDM): (a) Training sample probability distribution of facies 1; (b, c) Randomly selected training facies structures; (d) Reconstructed facies probability distribution; (e, f) LDM-generated facies structures; (g–i) Statistical comparisons of volume proportion, horizontal length, and vertical length; (j) Permeability distribution comparison; (k, l) Randomly generated permeability fields.

probability distribution that closely matches the distribution in the training samples. This reflects the ability of the LDM model to generate permeability fields that are consistent with facies types in a reasonable manner.

Although some existing works, such as Cui et al. (2024), achieve similar functionality by using two separate GANs to generate facies and permeability fields, these methods require different model structures and parameters for heterogeneous structure generation at different scales. This necessitates selecting and tuning models for each scale individually. In contrast, the LDM model achieves this within the same framework and hyperparameters, significantly simplifying the workflow for multi-scale aquifer heterogeneity modeling. This makes the LDM framework a promising tool for applications that require both large-scale structural features and fine-scale variations, such as pollutant transport modeling or multi-scale aquifer heterogeneity characterization.

4.3. Surrogate Modeling of Groundwater Flow and Solute Transport

After training, the groundwater flow and solute transport surrogate model based on the LDM framework, similar to other DL-based surrogate models, can quickly predict concentration and head distribution fields at multiple time steps by inputting different aquifer heterogeneity structures (in this case, facies structures). The prediction typically takes only a few seconds. In terms of simulation accuracy, Figures 3a and 3b show scatter plots comparing the true values from TOUGHREACT simulations with the concentration and head predictions for 1,000 randomly generated aquifer structures by the LDM model. The correlation coefficients for both head and concentration predictions exceed 0.99. Further spatial accuracy assessment of the surrogate model is conducted by comparing the predicted and true head and concentration fields. Figures 3d and 3e present results for two randomly selected aquifer structures, with predictions for both single time step head distributions and solute concentration fields at five different time steps. The predicted distributions (Figures 3d and 3e: Pred Con and Pred Head) closely match the distributions generated by TOUGHREACT (Figures 3d and 3e: True Con and True Head), highlighting the potential of the LDM framework in providing high-fidelity approximations of flow and transport fields.

In addition to facies-based predictions, the LDM surrogate model can be easily adapted for other representations of small-scale heterogeneity, such as permeability fields. By replacing the VAE training samples with permeability fields and adjusting the DDPM's conditional constraints to reflect these fields, the LDM can map permeability distributions directly to flow and transport outcomes. This flexibility makes the LDM a powerful alternative to traditional surrogate modeling approaches.

4.4. Direct Inversion of Aquifer Heterogeneity Structure

The direct inversion task is a challenging application that requires the model to infer large-scale facies structures directly from observed hydraulic head and solute concentration data. This task typically depends on the ability of DL models to capture the highly nonlinear relationships between observations and subsurface structures. For instance, Sun (2018) successfully applied two GANs to map hydraulic head fields to permeability fields. In this study, we investigate whether the same LDM framework used for the previous three tasks can achieve reliable identification accuracy in directly identifying aquifer heterogeneity.

Figures 3f and 3g illustrate examples where the LDM was applied to reconstruct facies distributions from two randomly selected sets of hydraulic head and solute concentration fields. The reconstructed facies closely match the true structures, demonstrating the framework's capacity for accurate inversion (Figures 3f and 3g: True facies structure and Predicted facies structure). Accuracy metrics in Figure 3c further validate the model's performance, with an average identification accuracy exceeding 92%. These results confirm that the LDM can effectively capture the complex relationships between hydraulic and solute transport observations and facies heterogeneity, thus realizing direct inverse identification of aquifer facies structure based on head and solute transport distribution field data.

This framework can also be extended to directly invert permeability fields from hydraulic head and solute concentration data. By replacing the VAE training samples with permeability fields and using permeability distributions as the conditioning constraints during DDPM training, the same LDM architecture can perform permeability field inversion.

Unlike traditional inversion methods, which rely on iterative numerical simulations and often incur significant computational costs, the LDM framework leverages its latent space to directly map observational data to

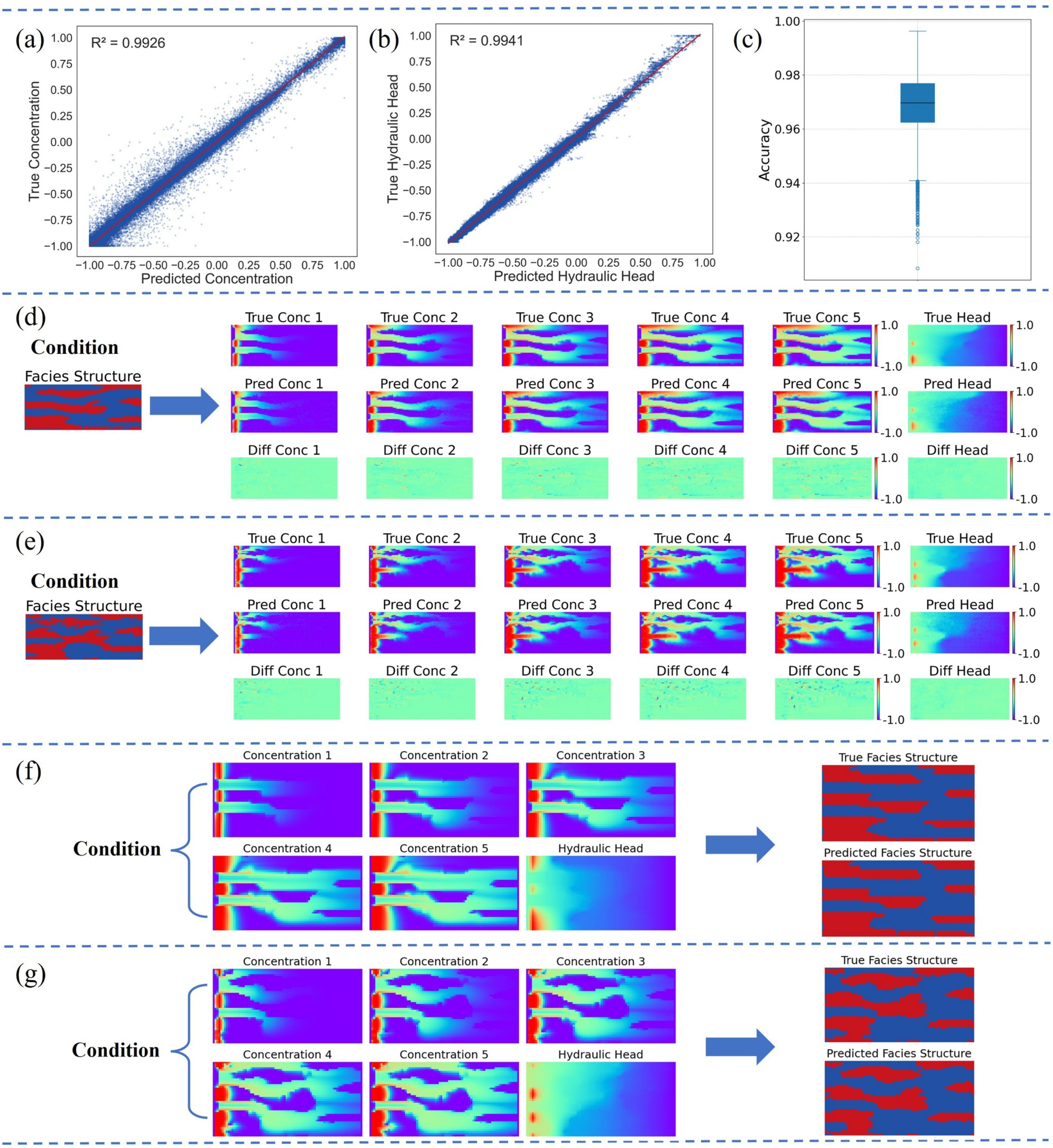

Figure 3. Performance of the Latent Diffusion Model (LDM) for surrogate modeling and aquifer structure inversion. Scatter plots (a, b) illustrate the comparison between predicted and true values for concentration and hydraulic head, respectively, (c) presents a box plot of aquifer structure identification accuracy. (d, e) Compare the predicted and true concentration and hydraulic head fields for two randomly selected aquifer facies structures, True Conc and True Head represent TOUGHREACT-predicted concentration and hydraulic head, while Pred Conc and True Head represent LDM-predicted concentration and hydraulic head. Diff Conc and Diff Head show the difference between the predicted and true concentration and hydraulic head, respectively. (f, g) Showcase the direct inversion of aquifer facies structures based on two randomly selected sets of concentration and hydraulic head fields.

subsurface heterogeneity. This eliminates the need for time-consuming iterative processes, significantly reducing computational overhead (Wang et al., 2024). Moreover, the simplicity and training stability of the LDM architecture make it a practical and robust alternative to existing DL-based inversion methods.

The LDM framework was implemented on a workstation with RTX4070 GPU and i9-14900K CPU, requiring approximately 2 hr for training and 1 s for inference per task. While these computational requirements are comparable to existing approaches, our framework uniquely handles multiple tasks using a single architecture and parameter set. Assuming traditional deep learning approaches require around 10 rounds of model architecture and parameter adjustments per task, our unified LDM framework could potentially reduce the development time to one-tenth by requiring only a single training process while achieving similar performance across all four tasks.

5. Discussion and Conclusions

This study develops a LDM framework capable of addressing multiple hydrogeological modeling tasks, including single-scale and multiscale aquifer heterogeneity generation, surrogate modeling for groundwater flow and solute transport, and direct inversion of aquifer heterogeneity. By employing a consistent architecture with shared hyperparameters, the LDM framework demonstrates strong generalization capabilities, effectively handling multiple objectives without the need for task-specific optimizations.

Despite these advancements, this work is limited to image-based conditional constraints in hydrogeological modeling. Future research could explore the incorporation of diverse data forms, such as text-based labels or point-based constraints, to broaden the range of applicable tasks. For example, textual labels like "stationary" or "non-stationary" could guide the generation of heterogeneity structures under specific statistical assumptions. Similarly, point-based constraints, such as borehole data or localized permeability measurements, could be integrated into the LDM training process to refine predictions further. Additionally, embedding physics-based constraints into the LDM framework, similar to Physics-Informed Neural Networks (PINNs), offers a promising avenue for ensuring physical consistency in generated outputs (Bastek et al., 2024; Shu et al., 2023).

Beyond image-based applications, the LDM framework also shows promise in tasks involving sequential or temporal data, such as groundwater level prediction. Its demonstrated strengths in video synthesis and sequence modeling (Blattmann et al., 2023; Gong et al., 2022) suggest that LDM could be extended to handle dynamic hydrogeological processes, providing accurate and efficient predictions over time.

In conclusion, while this study presents a novel approach for multi-tasking within hydrogeological modeling, demonstrating a step toward AGI, important limitations remain. Our model's ability to handle various tasks lays the groundwork for more generalized intelligent systems, but it is still constrained to specific, predefined hydrogeological tasks. Unlike true AGI systems, which can apply learned knowledge to unforeseen tasks and demonstrate autonomous reasoning across different domains, our framework operates within a narrow context. Nevertheless, the LDM framework's ability to generalize across diverse objectives using a consistent architecture highlights its promise as a foundational approach for future AGI systems in hydrogeology. While the current applications highlight the framework's strengths in handling image-based inputs, its adaptability to other data modalities and potential for incorporating diverse constraints and physical principles present opportunities for expanding its utility. This work contributes to the development of multi-tasking algorithms, but further research is required to achieve the broader, more flexible intelligence necessary for true AGI. By addressing existing limitations and exploring new applications, the LDM framework has the potential to redefine subsurface characterization and modeling practices, advancing hydrogeological research and applications.

Data Availability Statement

The codes are based on the implementations provided in the following repositories https://github.com/huggingface/diffusers and https://github.com/explainingai-code/StableDiffusion-PyTorch.

Acknowledgments
This work was funded by the National Natural Science Foundation of China (NSFC: U2267217, 42141011), Shandong Key Water Conservancy Science and Technology Project (2024370203001957) and Shandong Provincial Natural Science Foundation (ZR2024QD095).

References

Bastek, J.-H., Sun, W., & Kochmann, D. M. (2024). Physics-informed diffusion models. arXiv preprint arXiv:2403.14404.

Bengesi, S., El-Sayed, H., Sarker, M. K., Houkpati, Y., Irungu, J., & Oladunni, T. (2024). Advancements in generative AI: A comprehensive review of GANs, GPT, autoencoders, diffusion model, and transformers. *IEEE Access*, *12*, 69812–69837. https://doi.org/10.1109/ACCESS.2024.3397775

Blattmann, A., Rombach, R., Ling, H., Dockhorn, T., Kim, S. W., Fidler, S., & Kreis, K. (2023). Align your latents: High-resolution video synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 22563–22575). https://doi.org/10.1109/cvpr52729.2023.02161

Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., & Slooten, L. J. (2005). Inverse problem in hydrogeology. *Hydrogeology Journal*, *13*(1), 206–222. https://doi.org/10.1007/s10040-004-0404-7

Cui, Z., Chen, Q., Liu, G., & Xun, L. (2024). SA-RelayGANs: A novel framework for the characterization of complex hydrological structures based on GANs and self-attention mechanism. *Water Resources Research*, *60*(1), e2023WR035932. https://doi.org/10.1029/2023WR035932

Dai, Z., Ritzi, R. W., Jr., & Dominic, D. F. (2005). Improving permeability semivariograms with transition probability models of hierarchical sedimentary architecture derived from outcrop analog studies. *Water Resources Research*, *41*(7). https://doi.org/10.1029/2004WR003515

Di Federico, G., & Durlofsky, L. J. (2024). Latent diffusion models for parameterization and data assimilation of facies-based geomodels. arXiv preprint arXiv:2406.14815.

Ershadnia, R., Moeini, F., Abolfazl Hosseini, S., Dai, Z., & Reza Soltanian, M. (2024). Predicting multiphase flow behavior of methane in shallow unconfined aquifers using conditional deep convolutional generative adversarial network. *Journal of Hydrology*, *641*, 131674. https://doi.org/10.1016/j.jhydrol.2024.131674

Fei, N., Lu, Z., Gao, Y., Yang, G., Huo, Y., Wen, J., et al. (2022). Towards artificial general intelligence via a multimodal foundation model. *Nature Communications*, *13*(1), 3094. https://doi.org/10.1038/s41467-022-30761-2

Feng, T., Jin, C., Liu, J., Zhu, K., Tu, H., Cheng, Z., et al. (2024). How far are we from AGI. arXiv preprint arXiv:2405.10313.

Gershenzon, N. I., Ritzi, R. W., Jr., Dominic, D. F., Soltanian, M., Mehnert, E., & Okwen, R. T. (2015). Influence of small-scale fluvial architecture on CO_2 trapping processes in deep brine reservoirs. *Water Resources Research*, *51*(10), 8240–8256. https://doi.org/10.1002/2015WR017638

Gong, S., Li, M., Feng, J., Wu, Z., & Kong, L. (2022). Diffuseq: Sequence to sequence text generation with diffusion models. arXiv preprint arXiv:2210.08933.

Harp, D. R., Dai, Z., Wolfsberg, A. V., Vrugt, J. A., Robinson, B. A., & Vesselinov, V. V. (2008). Aquifer structure identification using stochastic inversion. *Geophysical Research Letters*, *35*(8). https://doi.org/10.1029/2008GL033585

Jankovic, I., Maghrebi, M., Fiori, A., & Dagan, G. (2017). When good statistical models of aquifer heterogeneity go right: The impact of aquifer permeability structures on 3D flow and transport. *Advances in Water Resources*, *100*, 199–211. https://doi.org/10.1016/j.advwatres.2016.10.024

Kang, X., Kokkinaki, A., Kitanidis, P. K., Shi, X., Lee, J., Mo, S., & Wu, J. (2021). Hydrogeophysical characterization of nonstationary DNAPL source zones by integrating a convolutional variational autoencoder and ensemble smoother. *Water Resources Research*, *57*(2), e2020WR028538. https://doi.org/10.1029/2020WR028538

Kitanidis, P. K. (2015). Persistent questions of heterogeneity, uncertainty, and scale in subsurface flow and transport. *Water Resources Research*, *51*(8), 5888–5904. https://doi.org/10.1002/2015WR017639

Kuusi, O., & Heinonen, S. (2022). Scenarios from artificial narrow intelligence to artificial general intelligence—Reviewing the results of the international work/technology 2050 study. *World Futures Review*, *14*(1), 65–79. https://doi.org/10.1177/19467567221101637

Laloy, E., Hérault, R., Jacques, D., & Linde, N. (2018). Training-image based geostatistical inversion using a spatial generative adversarial neural network. *Water Resources Research*, *54*(1), 381–406. https://doi.org/10.1002/2017WR022148

Lee, D., Ovanger, O., Eidsvik, J., Aune, E., Skauvold, J., & Hauge, R. (2024). Latent diffusion model for conditional reservoir facies generation. *Computers & Geosciences*, *194*, 105750. https://doi.org/10.1016/j.cageo.2024.105750

Mo, S., Zabaras, N., Shi, X., & Wu, J. (2020). Integration of adversarial autoencoders with residual dense convolutional networks for estimation of non-Gaussian hydraulic conductivities. *Water Resources Research*, *56*(2), e2019WR026082. https://doi.org/10.1029/2019WR026082

Moeini, F., Ershadnia, R., Rubinstein, R. L., Versteeg, R., Li, P., McGarr, J. T., et al. (2024). Employing generative adversarial neural networks as surrogate model for reactive transport modeling in the hyporheic zone. *Journal of Hydrology*, *639*, 131485. https://doi.org/10.1016/j.jhydrol.2024.131485

Rajaram, H., & Gelhar, L. W. (1995). Plume-Scale dependent dispersion in aquifers with a wide range of scales of heterogeneity. *Water Resources Research*, *31*(10), 2469–2482. https://doi.org/10.1029/95WR01723

Rizzo, C. B., & de Barros, F. P. J. (2019). Minimum hydraulic resistance uncertainty and the development of a connectivity-based iterative sampling strategy. *Water Resources Research*, *55*(7), 5593–5611. https://doi.org/10.1029/2019WR025269

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10674–10685). https://doi.org/10.1109/cvpr52688.2022.01042

Sbai, M. A. (2020). Unstructured gridding for MODFLOW from prior groundwater flow models: A new paradigm. *Groundwater Series*, *58*(5), 685–691. https://doi.org/10.1111/gwat.13025

Scheibe, T. D., Murphy, E. M., Chen, X., Rice, A. K., Carroll, K. C., Palmer, B. J., et al. (2015). An analysis platform for multiscale hydrogeologic modeling with emphasis on hybrid multiscale methods. *Groundwater Series*, *53*(1), 38–56. https://doi.org/10.1111/gwat.12179

Shu, D., Li, Z., & Barati Farimani, A. (2023). A physics-informed diffusion model for high-fidelity flow field reconstruction. *Journal of Computational Physics*, *478*, 111972. https://doi.org/10.1016/j.jcp.2023.111972

Song, X., Chen, X., Ye, M., Dai, Z., Hammond, G., & Zachara, J. M. (2019). Delineating facies spatial distribution by integrating ensemble data assimilationand indicator geostatistics with level-set transformation. *Water Resources Research*, *55*(4), 2652–2671. https://doi.org/10.1029/2018WR023262

Sun, A. Y. (2018). Discovering state-parameter mappings in subsurface models using generative adversarial networks. *Geophysical Research Letters*, *45*(20), 11137–11146. https://doi.org/10.1029/2018GL080404

Wang, N., Kong, X.-Z., & Zhang, D. (2024). Physics-Informed Convolutional Decoder (PICD): A novel approach for direct inversion of heterogeneous subsurface flow. *Geophysical Research Letters*, *51*(13), e2024GL108163. https://doi.org/10.1029/2024GL108163

Yang, J., Zhang, D., & Lu, Z. (2004). Stochastic analysis of saturated–unsaturated flow in heterogeneous media by combining Karhunen-Loeve expansion and perturbation method. *Journal of Hydrology*, *294*(1), 18–38. https://doi.org/10.1016/j.jhydrol.2003.10.023

Yang, L., Zhang, Z., Song, Y., Hong, S., Xu, R., Zhao, Y., et al. (2023). Diffusion models: A comprehensive survey of methods and applications. *ACM Computing Surveys*, *56*(4), 1–39. Article 105. https://doi.org/10.1145/3626235

Zhan, C., Dai, Z., Soltanian, M. R., & Zhang, X. (2022). Stage-wise stochastic deep learning inversion framework for subsurface sedimentary structure identification. *Geophysical Research Letters*, *49*(1), e2021GL095823. https://doi.org/10.1029/2021GL095823

Zhan, C., Dai, Z., Yang, Z., Zhang, X., Ma, Z., Thanh, H. V., & Soltanian, M. R. (2023). Subsurface sedimentary structure identification using deep learning: A review. *Earth-Science Reviews*, *239*, 104370. https://doi.org/10.1016/j.earscirev.2023.104370

Zhang, J., Cao, C., Nan, T., Ju, L., Zhou, H., & Zeng, L. (2024). A novel deep learning approach for data assimilation of complex hydrological systems. *Water Resources Research*, *60*(2), e2023WR035389. https://doi.org/10.1029/2023WR035389

Zhang, X., Jiang, S., Wei, J., Wu, C., Xia, X., Wang, X., et al. (2024). Non-Gaussian hydraulic conductivity and potential contaminant source identification: A comparison of two advanced DLPM-based inversion framework. *Journal of Hydrology*, *638*, 131540. https://doi.org/10.1016/j.jhydrol.2024.131540

Zhang, X., Jiang, S., Zheng, N., Xia, X., Li, Z., Zhang, R., et al. (2024). Integration of DDPM and ILUES for simultaneous identification of contaminant source parameters and non-Gaussian channelized hydraulic conductivity field. *Water Resources Research*, *60*(9), e2023WR036893. https://doi.org/10.1029/2023WR036893

Zhu, J., & Yeh, T.-C. J. (2005). Characterization of aquifer heterogeneity using transient hydraulic tomography. *Water Resources Research*, *41*(7). https://doi.org/10.1029/2004WR003790

References From the Supporting Information

Carle, S. F., & Fogg, G. E. (1996). Transition probability-based indicator geostatistics. *Mathematical Geology*, *28*(4), 453–476. https://doi.org/10.1007/BF02083656

Carle, S. F., & Fogg, G. E. (1997). Modeling spatial variability with one and multidimensional continuous-Lag Markov chains. *Mathematical Geology*, *29*(7), 891–918. https://doi.org/10.1023/A:1022303706942

Dai, Z., Middleton, R., Viswanathan, H., Fessenden-Rahn, J., Bauman, J., Pawar, R., et al. (2014). An integrated framework for optimizing CO_2 sequestration and enhanced oil recovery. *Environmental Science and Technology Letters*, *1*(1), 49–54. https://doi.org/10.1021/ez4001033

Dai, Z., Wolfsberg, A., Lu, Z., & Ritzi, R., Jr. (2007). Representing aquifer architecture in macrodispersivity models with an analytical solution of the transition probability matrix. *Geophysical Research Letters*, *34*(20). https://doi.org/10.1029/2007GL031608

Deutsch, C. V., & Journel, A. G. (1992). Geostatistical software library and user's guide. *New York*, *119*(147).

Zhan, C., Dai, Z., Samper, J., Yin, S., Ershadnia, R., Zhang, X., et al. (2022). An integrated inversion framework for heterogeneous aquifer structure identification with single-sample generative adversarial network. *Journal of Hydrology*, 127844. https://doi.org/10.1016/j.jhydrol.2022.127844